

Parametric Facial Expression Synthesis and Animation Review

Mohammad Amin Saeed, Behzad Mahjour Shafiei, Fariborz Iranmanesh

Kerman Branch, Islamic Azad University, Kerman, Iran.

Mohammad Amin Saeed, Behzad Mahjour Shafiei, Fariborz Iranmanesh; Parametric Facial Expression Synthesis and Animation Review

ABSTRACT

In this paper, we proposed a parametric model based facial animation synthesis. In the proposed algorithm various facial expressions are generated using the parametric models and the novelty of the algorithm is that new expressions can be generated by blending selected expressions. A complete parameterized model for faces would allow the animator to create any facial image by specifying the appropriate set of parameter values. Using the interface developed along with the proposed algorithm, these images can be used to develop an animation based on an interactive system, virtual reality and games. Interface allows selection of linear and non linear interpolation, play animation and save animation. In this paper we have highlighted advantages and disadvantages of linear and non linear interpolation and proposed a method which incorporates advantages of both.

Key words: Parametric model, facial animation, blending, interpolation, key-frame etc.

Introduction

The methods used for generating facial animation are mainly classified in three categories in the literature [17]: Key framing, Parameterization and Muscle-based modeling. In key frame animation [5], the desired images are specified at different points of time. A computer algorithm generates the necessary images for the frames in between these key frames. Parke and Waters [1] are pioneers in the facial modeling and animation. In [1], the key framing has been used and the face is animated using cosine interpolation. Developing the required 3-D key frame descriptions is difficult as the number of key frames needed becomes large. As a result new system was proposed which defines a set of expression parameters for different parts of the face [5]. Each expression parameter affects a set of vertices of the face model and generates key frame. The disadvantage of the parametric systems is that they cannot easily blend facial expressions since each parameter affects a disjoint set of vertices in the face model.

Platt S M [2] has proposed structured based facial models which are based on the anatomy of the face. However, all these models do not take into account the fact that the face is not only a geometric model but a complex biomechanical system. So Terzopoulos and Waters [3] have proposed physically based face models to overcome these limitations. They model the face in a layered fashion and incorporate an anatomically based muscle model

with a physically based layered tissue model [18]. Anatomically based muscle models [4] are first used by waters to animate major facial expressions. Later methods based on parameterization and muscle based modeling [5, 4] are proven to be efficient and provides the realistic view. The problem with these approaches is that the number of parameters is very large and the model is very complex because of muscle structure. In this paper, we proposed the hybrid technique for facial animation generation which uses key framing but the novelty is that key frames are described by the set of parameters.

One of the most important issues in the facial animation is representing the universal expressions of the face, such as fear, anger, happiness, surprise and sadness etc. In our proposed system these expressions are represented by the parametric models. The proposed algorithm needs only the set of values for these parameters like lip width, jaw angle, eyes width, eyelid position, eyebrow position, gaze height and gaze width. Based on these parameters the position of corresponding vertices is interpolated. We also propose a method to generate new expression through the blending of selected expressions. Users can select key frames and intermediate frames are generated by using linear and non linear interpolation. Realistic face image is produced by changing the position and color of light source. More realism is added with texture mapping.

The paper is organized as follows: Section 2 describes face model. It is followed by expression generation and facial animation in section 3 and 4

Corresponding Author

Mohammad Amin Saeed Computer software engineering student, Islamic Azad University, Kerman Branch, Kerman, Iran.
E-mail: amin.saeed.89@gmail.com

respectively. Section 5 describes graphical user interface. The simulation results and conclusion are discussed in section 6 and 7.

Proposed Face Model:

The face model is the part of head model. The head model may be obtained using photographs [10], 3D scanning [19] or modeling using special software. Models can use different representations, like polygon mesh and B-Spline. The animation can also be generated in several different ways, like morphing geometry between different pre-modeled faces, simulating muscles of the face, or using sensors or trackers. In the proposed method, we have created polygon-based (triangle mesh) head model to

generate realistic image of face. Wire-frame model is shown in Figure 1.

In addition to accurately capturing the shape of the head it is also important to structure its geometry in a way that will later permit naturalistic facial animation. This typically involves arranging the polygons of the surface to align with the muscles of the face. Notice in the wire-frame model that the polygons conform to the shape of the cheek, lip, eye, forehead, ears, nose and eye-brows etc. The polygons are sized and positioned to match the features of the face. There are many small polygons in areas of high curvature and fewer, larger polygons in the flatter regions. Also, polygon boundaries are designed to coincide with the desired creases in the face. Our proposed model consists of 350 triangles and 215 vertices.

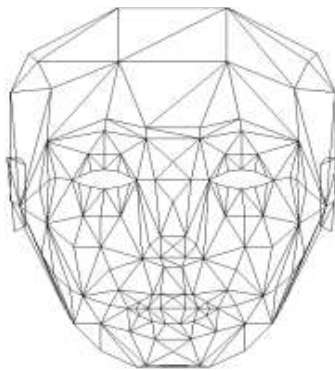


Fig. 1: 3D Mesh model of face.

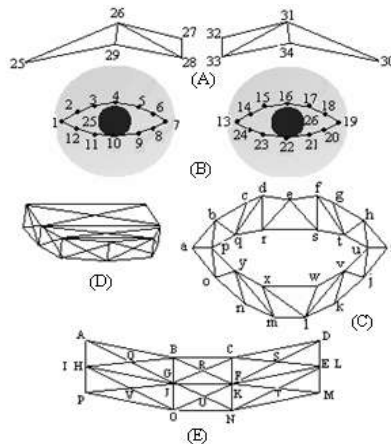


Fig. 2: (A) Eyebrow (B) Eyes (C) Mouth (D) Tongue (E) Teeth.

For this model, image generation is based on surfaces constructed from connected networks of polygons. The facial mask is one polygonal surface, and lip[6,14] is another; teeth, tongue and eyebrows are separate polygonal surfaces as shown in Figure 2. To model the eye, two spheres with different radius are used. Smaller sphere is drawn with black color to form pupil and big sphere is drawn with white color as shown in Figure 2.

The polygonal connection, or topology, of each surface depicted in Figure 1 and Figure 2 remains constant, but the three dimensional position of each polygon vertex varies according to the parameter values, eye orientation, and face orientation etc. As the vertex positions change, the polygonal surfaces flex and stretch to allow the face to change.



Fig. 3: Rendered face.

For simulation, we have rendered the face model using OpenGL [12]. In Figure 3, we see this model with smooth shading, realistic textures and lighting effects. Once the model has been built, any of the facial animation techniques can be used to animate it. This model is implemented in such a way that it is both data and parameter driven. Data files define the polygon topology and the extreme vertex positions used for interpolation. The parameter values can be selected interactively.

2. Expression Generation:

Parameterized models can produce realistic images of human faces with a surprisingly small number of parameters [5] and that can be manipulated easily. One of the most widely used techniques in facial animation is key-framing based on facial parameters. Combinations of parameters provide a large range of facial expressions. Expressions are generated by moving and deforming various control vertices of face model according to parameter values.

2.1 Expression Parameters:

Expression parameters found useful for the eyes include pupil dilation, eyelid opening and eyebrow position. Useful expression parameters for the mouth include jaw rotation (which controls mouth opening), width of the mouth, mouth expression (smiling, frowning, etc.), position of the upper lip, and positions of the corners of the mouth [21,22]. Other useful expression parameters include the orientation of face (head) with respect to the neck and the rest of the body. With very few parameters, it is possible to develop a model that allows interesting expression animation and can be animated to a spoken soundtrack.

Our system provides left/right head movement by rotating about y-axis, up/down head movement by rotating about x-axis and head tilt by rotating about z-axis. Various facial expressions are generated by changing the mouth width, mouth opening/closing, eye-brow position, eye opening/closing and changing the position of pupil. Table 1 summarizes all parameters that we have used and its corresponding vertices.

Table 1: Parameters and their corresponding vertices.

Face expression	parameter	vertices
Head up/Down	X angle	All
Head left/right	y angle	All
Head tilt left/right	Z angle	All
Mouth wide/pucker	Lip width	X position of a, b, c, d, f, g, h, I, j, h, k, l, m, n, o, p, q, r, s, t, u, v,x,y
Mouth open/ close	Jawangle	X position of a, i, j, k, l, m, n, o, p, u, v, w, x, y, J, K, O, N, T,U, V
Eye-brow up/Down	Eyebrow	y position of 25, 27, 28, 30,32, 33
Eye open/ close	Eyelid	y position of 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 14, 15, 16, 17, 18, 20, 21, 22, 23, 24
Pupil up/Down left/right	Gaze height Gaze width	X and y position of 25, 26
smile	smile	X and y positoin of a, b, p, o, h, I, j, u

2.2 Facial Expression:

We define an expression as any change in the face geometry. In our system user can generate any expression by selecting appropriate parameters

values. System also provides facial expression like natural, fear, sad, tired, surprise, angry and happy [20, 23]. In Table 2 various facial expressions of our system and responsible parameters are displayed.

Table 2: Facial Expressions and corresponding parameters.

expression	Parameters
Fear	Yangle, jawangle, eyebrow, eyelid
Sad	jawangle, eyebrow, eyelid
Tired	Yangle, Eyelid zangle
surprise	jawangle, eyebrow, eyelid
angry	jawangle, eyebrow, eyelid
happy	jawangle, Eyelid, lipwidht, smile

2.3 Blend Shape Technique:

Shape interpolation (blend shapes, morph targets, and shape interpolation) is the most intuitive and commonly used technique in facial animation practice. A blend shape model [7] is simply the linear weighted sum of a number of topologically conforming shape primitives. It is given as (1).

$$V_j = \sum W_k * B_{kj} \tag{1}$$

V_j is the j^{th} vertex of resulting model, W_k is blending weight ($0 < W_k < 1$), and B_{kj} is the j^{th} vertex of the k^{th} blend shape. The weighted sum can be applied to the vertices of polygonal models. This technique combines the effects of multiple

expressions. Figure 4 shows an example of result of blending two expressions. Simpler expressions can be combined visually and interactively to create more complex expressions.

2.4 Rendering:

Smooth shading [16] is used to form realistic face image. More realism is added with texture mapping [13] as shown in Figure 5. In this expression, we have mapped texture image of scratches (see Figure 6a) on forehead. For realistic look of teeth we have mapped texture image of teeth (see Figure 6b) to teeth surface. User can also change the position of the light source.



Fig. 4: Generating expressions with blending shape.



Fig. 5: Achieving reality with texture mapping.



Fig. 6: (a) Scratches image (b) Teeth image.

3. Animation Generation:

Once the expressions are modeled as described above, it is time to animate our character [15] of interest. As mentioned earlier, we have used key-frame approach of animation. Every facial expression can be stored as key-frame by storing the values of parameters. After all key-frames are defined, animation can be created by generating intermediate frames. Intermediate frames are generated by interpolating the parameters values of successive key-frames. Our algorithm supports three types of interpolation: (i) Linear, (ii) Cubic B-spline and (iii) Cardinal Spline. When number of key frames is more than two, with linear interpolation jerky and abrupt movement is observed at joint point of key-frames [8]. Smooth animation is achieved with the help of non-linear animation (B-spline and Cardinal Spline). Advantage of Cardinal spline over B-spline is that the curve interpolates all the points specified. In our developed interface, when user select to save the animation, only parameters values of all key-frames are saved in the file. The same animation can be loaded using 'Load Animation' option and 'Play Animation' option allows us to play a loaded animation. This following section describes the different interpolation techniques in brief.

3.1 Linear Interpolation:

Linear interpolation between two points $p[i]$ and $p[i+1]$ is given by (2). It results in straight line between two points.

$$\text{Parameter} = P[i] + f * (P[i+1] - P[i]) / (N-1), \quad 0 \leq f \leq N-1 \quad (2)$$

Here f is current frame and N is total number of frames between two key frames.

3.2 Cubic B-Spline Interpolation:

In case of B-spline [9] and Cardinal Curves [8], the curves are generated considering the value of each parameter of each key frame as control vertex. Cubic B-spline curve starting nearby $p[i]$ and end nearby $p[i+1]$ is defined as

$$Q(F) = ((1-F)^3/6) * p[i-1] + ((3 * F^3 - 6 * F^2 + 4) / 6) * p[i] + ((-3 * F^3 + 3 * F^2 + 3 * F + 1) / 6) * p[i+1] +$$

$$(F^3/6) * p[i+2] \quad (3)$$

$F = 1.0 * f / (N-1) \quad 0 \leq f \leq N-1$ where f indicate current frame and N indicate total number of frames between two key frames.

3.3 Cardinal Spline Interpolation:

A cardinal spline section is specified with four consecutive control points in which two middle control points are section endpoints and other two are used in the calculation of endpoint tangents. Special care is taken so curve interpolates first and last point properly. Cardinal spline passing through points $p[i]$ and $p[i+1]$ is specified with four points $p[i-1], p[i], p[i+1]$ and $p[i+2]$. Tension parameter controls looseness of the cardinal spline section. Smaller values of tension tend to produce tighter curves and the motion at the control point is more abrupt. Increasing the value of tension will relax the curve at the control points and slow down the motion at the control points. Figure 7 give the animation plot for tension=0.1 and tension=1.0.

$$\text{Tangent1}(Q'(p[i])) = (1 - \text{tension}) * (p[i+1] - p[i-1]) / 2 \quad (4)$$

$$\text{Tangent2}(Q'(p[i+1])) = (1 - \text{tension}) * (p[i+2] - p[i]) / 2 \quad (5)$$

$$Q(F) = (2 * F^3 - 3 * F^2 + 1) * p[i] + (-2 * F^3 + 3 * F^2) * p[i+1] + (F^3 - 2 * F^2 + F) * \text{tangent1} + (F^3 - F^2) * \text{tangent2} \quad (6)$$

$F = 1.0 * f / (N-1) \quad 0 \leq f \leq N-1$ where f indicate current frame and N indicate total number of frames between two key frames.

4. Graphical User Interface:

With the proposed algorithm, complete animation system has a graphical user interface as shown in Figure 8 developed using GLUT user interface library [11]. User interface is event-driven and contains buttons to draw the face according to the specified rendering mode like wireframe, flat/smooth shaded and textured. Also GUI support buttons for eye control, mouth control, head control and various expressions. Buttons are also provided to set key-frame, save animation, play animation, save picture (image of face in BMP format [8]), select type of interpolation, change the position and color of light source and selecting number of frames per second.

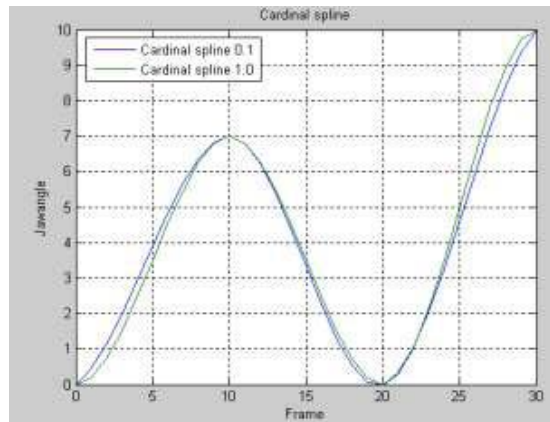


Fig. 7: Animation plot for tension=0.1 and 1.0.

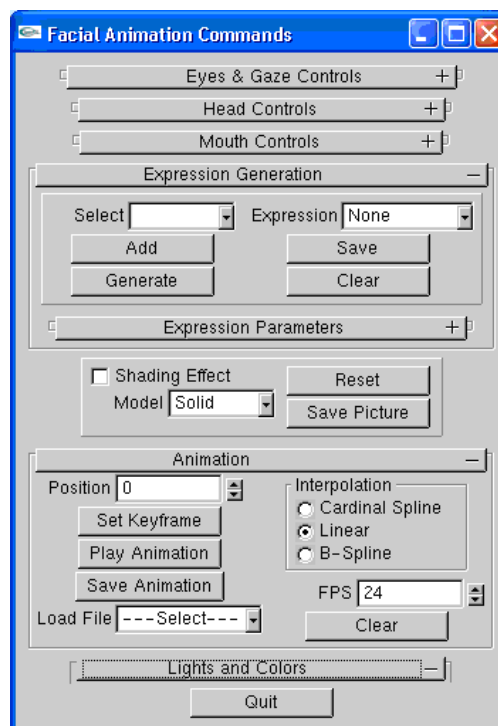


Fig. 8: Graphical user interface.

Simulation and Results

The proposed algorithm is implemented using OpenGL and GUI is designed as discussed above. Animation frames generated by Linear Interpolation are shown Figure in 9. The faces with different expressions like fear, sad, tired, surprise, angry and happy is shown in Figure 10. The facial states could be saved as a key-frame together with a light position. Figure 11 depicts various expressions generated by applying blending technique to combine expressions presented in Figure 10. In Figure 12, face rendered by changing diffuse color,

specular color and light position is shown. Figure 13 shows the comparison among three interpolation techniques. When key-frames are more than two non-linear interpolations give better result than linear interpolation. Non-linear interpolation does not retain values of adjacent key frames when these are equal. Figure 14 shows that jaw angle value remains same during key frame 10 and 20 but Cardinal spline does not retain value. So we have modified Cardinal spline interpolation which dynamically make decision and do not do interpolation between Key-frames whose parameters value remains same. Figure 15 shows the correct result for same.

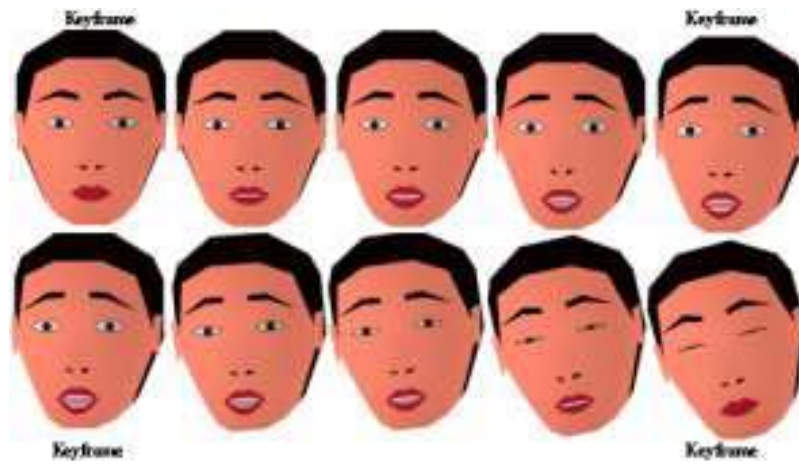


Fig. 9: Animation frames generated using linear interpolation.

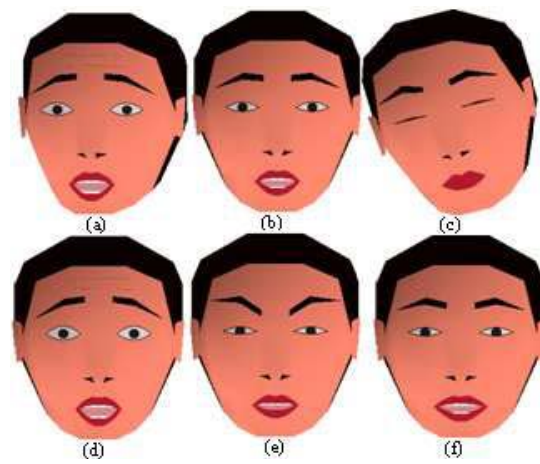


Fig. 10: Expressions (a) Fear, (b) Sad, (c) Tired, (d) Surprise, (e) Angry and (f) Happy.

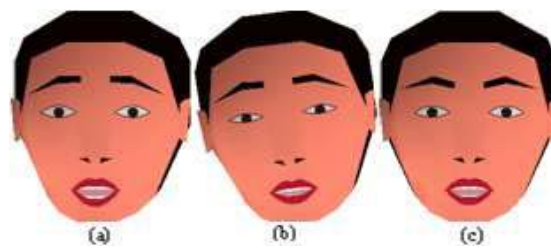


Fig. 11: Expressions generated with Blending (a) Fear and Sad, (b) Tired and Surprise and (c) Angry and Surprise.

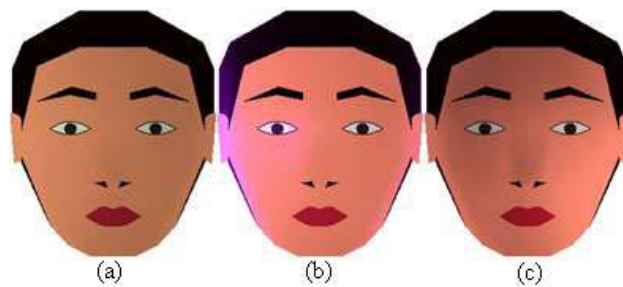


Fig. 12: Face rendered by changing (a) Diffuse Color (b) Specular Color (c) Light Position.

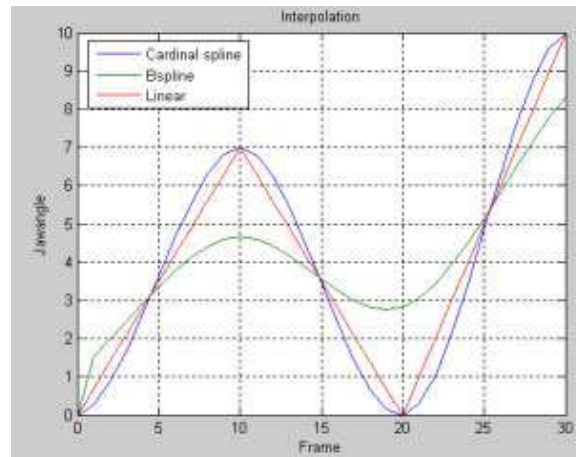


Fig. 13: Comparisons among three interpolation techniques.

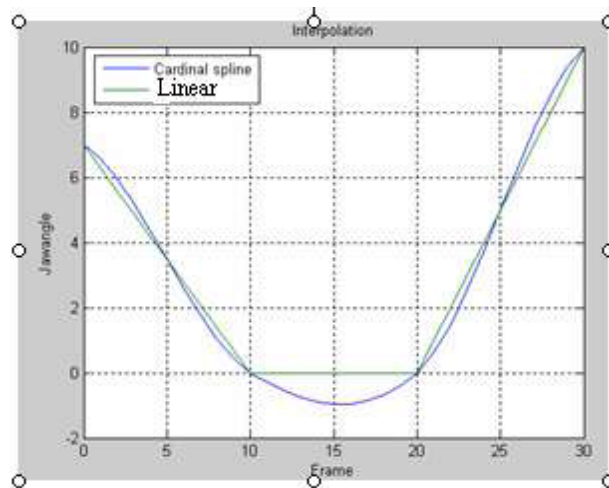


Fig. 14: Comparisons between linear and cardinal spline Interpolation

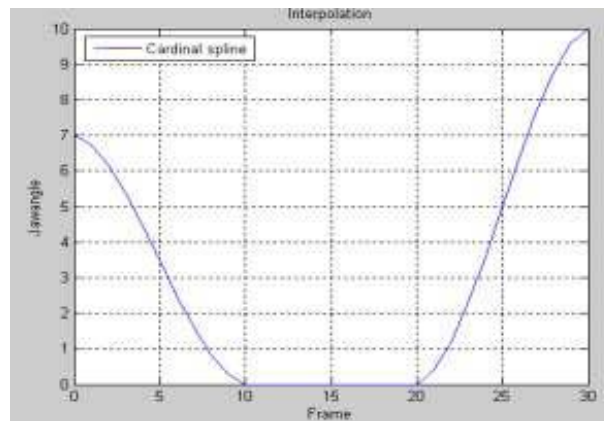


Fig. 15: Modified cardinal spline interpolation.

Conclusion:

Our proposed algorithm allows one to define the parametric models for various expressions and it also generates new expression by applying the blending to

the existing expressions. The interactive system developed as a part of the proposed algorithm is very useful tool to design the facial animation. Our parametric model for expressions allows us to define any expression in terms of few parameters and using

the interpolation the sequence of images can be generated for animation. User is also able to work with solid model as well as wire-frame model. Face with single or multiple expressions can be saved as key-frame. User can see the effect of linear, Cubic B-spline and Cardinal spline interpolation during playing animation. Cardinal spline interpolation is giving best result even though parameters retain values between two adjust key frames. User can save entire animation and play it later on. Our system is very useful for synthesis realistic face by changing position of light source and material properties like diffuse and specular color.

Our system provides all features necessary for complete facial animation system. Existing systems do not provide all of them. Expression generation by blending technique, texture mapping for realistic look and user friendly Graphical User Interface are distinctive features of our system. The proposed algorithm can be extended to generate facial animation by incorporating sound with facial expressions.

References

1. Parke, F.I., 1972. Computer generated animation of faces. Proc of ACM National conference., 1: 451-457.
2. Platt, S.M., 1985. A structural model of the human face. Ph.D. Thesis, University of Pennsylvania dept of CIS.
3. Terzopoulos, D. and K. Waters, 1990. Physically based facial modeling, analysis and animation. The journal of visualization and computer animation., 1(4): 73- 80.
4. Waters, K., 1987. A muscle model for animating 3D facial expression. ACM computer graphics, 21(4): 17- 24.
5. Parke, F.I., 1982. Parameterized models for facial animation. IEEE computer graphics and applications. 2(9): 61-68.
6. Ze-jing chuang and Wu. Chung-hsien, 2002. Parameter based Lip modeling for facial animation of general objects. In Proceedings of ICME., 1: 453-456.
7. Zhigang Deng and Ulrich Neumann, 2008. Data-Driven 3D Facial Animation. Springer Publication.
8. Shalini Govil-pai, 2007. Principles of computer graphics theory and practice using OpenGL and Maya. First edition, springer.
9. David, F Rogers and J Alal Adams, 2002. Mathematical elements for computer graphics, second edition, Tata McGraw-Hill.
10. Pighin, F., J. Hecker, D. Lischinski, R. szeliski and D.H. salesin, 1998. Synthesizing realistic facial expressions from photographs, In Proceedings of SIGGRAPH. pp: 75-84.
11. Paul Rademacher, 1999. GLUI: A GLUT-Based User Interface Library. Version 2.0.
12. Mason, Woo., 1997. OpenGL Architectural Review Board, Jackie Neider, Tom Davis, OpenGL programming guide. Second edition.
13. Rocchini, C., P. Cignoni, C. Montani and R. scopigno, 1999. Multiple textures stitching and blending 3D objects in rendering techniques. Eurographics workshop on rendering, pp: 119-130.
14. Thierry Guiard-Marigny, Nicolas Tsingos, Ali adjoudani, Christian Benoit and Marie-paule Gascuel, 1996. 3D models of the Lips for realistic speech animation. In proceedings of the computer animation apos., pp: 80-89.
15. Platt S.M. and N.I. Badler, 1981. Animating Facial Expressions. ACM SIGGRAPH computer graphics, 15(3): 245-252.
16. Phong. B.T., 1976. Illumination for computer pictures. CACM., 18(6): 311-317.
17. Arai, K., T. Kurihara and K. Anjyo, 1996. Bilinear interpolation for facial expression and metamorphosis in real time animation. Journal of the visual computer., 12: 105-116.
18. Lee, Y., D. Terzopoulos and K. Waters, 1993. Constructing physical-based facial models of Individuals. In Proceedings of Graphics Interface '93. Canadian Information, Processing Society.
19. Lee, Y., D. Terzopoulos and K. Waters, 1995. Realistic modeling for facial animation. Computer Graphics. 29(4): 55-62.
20. Jeong Ho Lee, Jong Min Lee, Hyung, Jin Kim and Young Shik Moon, 2007. Automatic synthesis of realistic facial expressions. IEEE international symposium on signal processing. pp: 46-51.
21. Yongmian, Zhang, Qiang, Ji, Zhiwei and yi. Beifang, 2008. Dynamic facial expression analysis and synthesis with MPEG-4 facial animation parameters. IEEE transactions on circuits and systems for video technology, 18(10).
22. Malate, L., A. Raouzaoui, K. Karpouzis and S. Kollias, 2009. MPEG-4 Facial expression synthesis. Journal of personal and ubiquitous computing., 13: 77-83.
23. Wang, hao., 2009. Facial expression synthesis and analysis. Journal of E-business and Telecommunications. 23: 269.